

CERES Data Management Working Group Status Report

CERES Science Team Meeting

April 22, 2014

Jonathan Gleason
Data Management Lead



Overview



- Software Status Update
- Recent DMT Activity
- Current Product Availability
- Production Automation Activity (CATALYST Status)
- Edition 3 & Edition 4 Processing Approach
- CERES Data User Statistics



Data Management Team

**Systems/
Optimization:**
Nelson Hillyer
Josh Wilkins

**Production /
Optimization:**
Lisa Coleman
Carla Grune

Instrument:
Denise Cooper
Thomas Grepiotis
Dianne Snyder

Convolution:
Igor Antropov

Inversion:
Victor Sothcott

**Product Web
Support:**
Churngwei Chu
Ed Kizer
Cristian Mitrescu

**Configuration
Management:**
Tammy Ayers
Joanne Saunders

ERBELike:
Dale Walikainen
Jeremie Lande

SARB:
Tom Caldwell

Task Manager:
Walt Miller

Clouds:
Sunny Sun-Mack
Ricky Brown
Yan Chen
Liz Heckert
Rita Smith
Sharon Gibson

TISA Gridding:
Raja Raju
Forrest Wrenn

TISA Averaging:
Cathy Nguyen
Dennis Keyes

CERES Production Software Overview

Subsystem Number	Subsystem Name	Number of PGEs	LOC (to nearest 1K)	Publicly Available Data Products	Product Frequency
1	Instrument/Pre-Processor	1	9K		
1	Instrument	5 + lib	157K	BDS	1/day
2	ERBE-like/ Inversion	5 + lib	31K	ES-8	1/day
3	ERBE-like/ TSA	2	14K	ES-9, ES-4	1/month
4.1 – 4.4	Clouds/VIIRS Subset Code	1	21K		12/hour
4.1 – 4.4	Clouds	11	358K		
4.5 – 4.6	Inversion	11	110K	SSF	1/hour
5	SARB	1	36K		1/hour
6 & 9	TISA-Gridding	7	34K	SSF1deg-Hour, ISCCP-D2like-Day/Nit	60/month, 36/month, 1/month
11	GGEO	1	6K	ISCCP-D2like-GEO	1/month
7.2	Synoptic SARB	1	47K		
7.1 & 8 10	TISA-Averaging	4	102K	SSF1deg-Day, SSF1deg-Month, SYN1deg-(3Hour, M3hour, Mhour, Month)	1/day, 1/month, 1/month 5/month
12	MOA	2	14K		
	CERESlib/Perl_Lib		131K		
	Total	53	1,070K		



Recent Activity

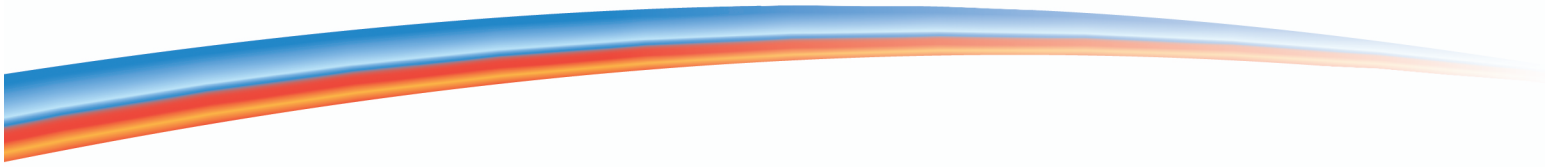


30 Software & Data Deliveries Since October 25th, 2013

- **CATALYST & PR Tool (10)**
 - Build 1 Server and Console
 - Logging DB Updates & incremental patch
 - CATALYST status monitoring utility
 - PR Tool CATALYST updates
- **Instrument (1)**
 - Time-varying gains for NPP-FM5 Edition1
- **Clouds (2)**
 - Migration of Clouds PGEs from *Magneto* to *AMI-P* & Edition1 NPP Code
- **Inversion (3)**
 - Ed4 Delivery, SIBI preprocessor, Remove SCOOL output files
- **SARB (3)**
 - Ancillary input files
- **TISA Gridding (4)**
 - Ed3 ISCCP-D2like-Day/Nit & ISCCP-D2like-Mrg supporting documentation
- **TISA Averaging (1)**
 - SOURCE ancillary input file for data through December 2013
- **GGEO (1)**
 - ISCCP-D2like-GEO supporting documentation
- **Perl_Lib (4)**
 - CATALYST updates
- **CERESlib (1)**
 - New memory request env variable for x86.

Current Data Availability

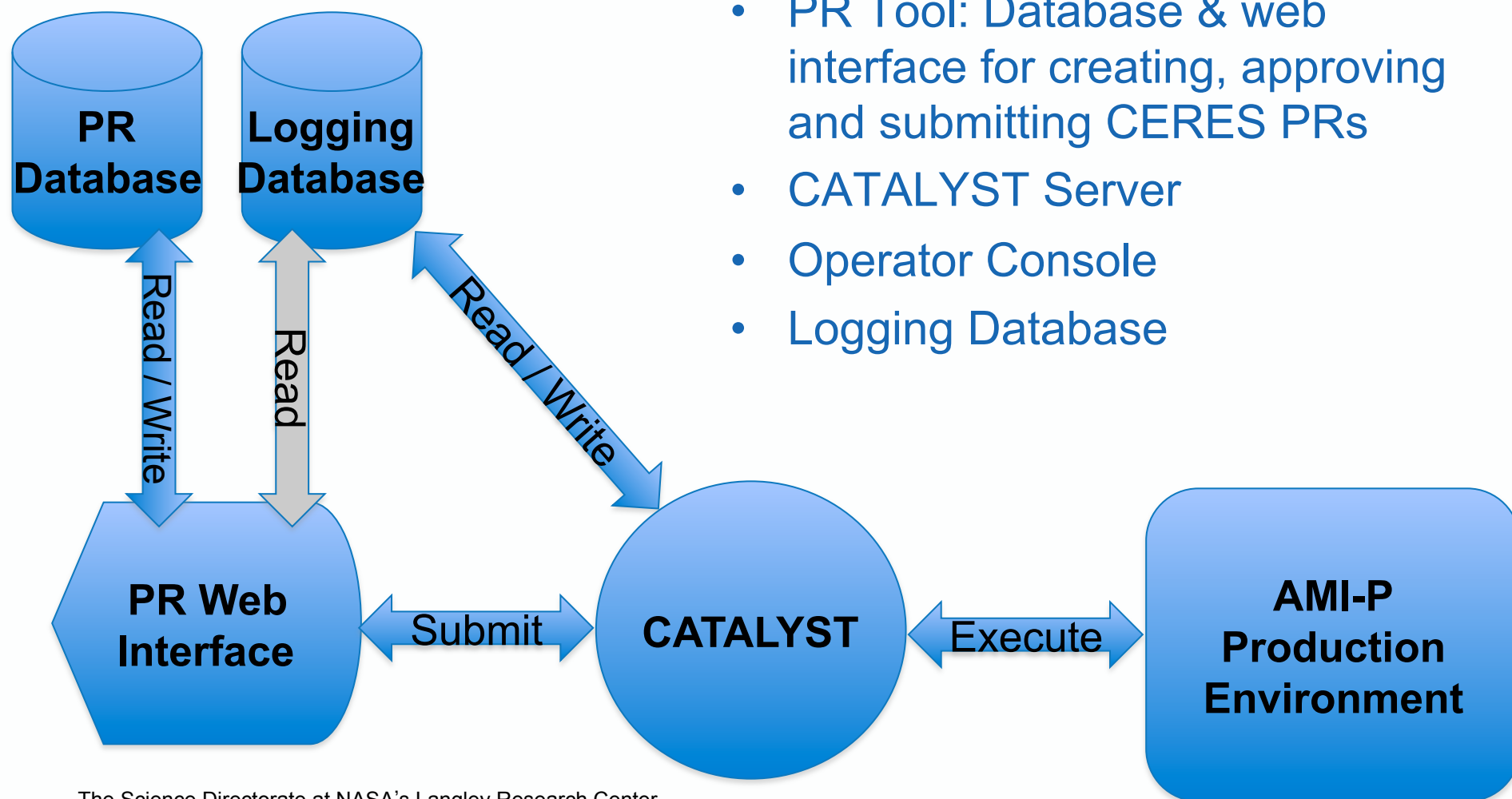
Product	Platform	Processed through	Publically Available
BDS (Ed 3)	Terra & Aqua	Dec 31, 2013	Yes
SSF (Ed 3)	Terra & Aqua	Dec 31, 2013 (FM1 & FM3 only)	Yes
SFC (Ed 3)	Terra & Aqua	Nov 2013	Yes
SSF1deg-Month (Ed 3) SSF1deg-Day	Terra & Aqua	Sept 2013	No
SYN1deg (Ed 3)	Merged	Nov 2013	Yes
ISCCP-D2like-Day/Nit (Ed3)	Terra Aqua	Dec 2013 Oct 2013	Yes
ISCCP-D2like-GEO (Ed3)	GEO	June 2013	Yes
ISCCP-D2like-Mrg (Ed3)	Terra + Aqua + GEO	Dec 2013	Yes
SSF Beta2 Ed4	Terra Aqua	June 1, 2005 Mar 31, 2005	No
SSF Edition 1-CV	Terra Aqua	Feb 28, 2014	No



Production Automation Effort



Automation Framework

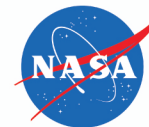




CATALYST



- **CERES AuTomAteD job Loading sYSTem (CATALYST)**
- Automation Framework for PGE Execution, Coordination & Logging that:
- Modularly implements science software dependency logic
- Leverages existing CERES production infrastructure
 - Ingest CERES PR to create collection of jobs
 - Build all possible individual jobs
 - Identify job input dependencies
 - Execute on AMI-P cluster
 - Send jobs to Sun Grid Engine
 - Utilize existing job submission scripts
 - Initiate ANGe ingest wrapper scripts
- Provides graphical interface for users to manage system



CATALYST (Cont'd)

- Automation Framework for PGE Execution, Coordination & Logging that:
 - Dynamically manages job input dependencies and determine when jobs are runnable
 - CATALYST jobs wait for predecessor jobs to complete
 - Predecessor jobs can be:
 - Internal to CATALYST
 - External to CATALYST (via a backloging interface)
 - Jobs broadcast completion status to follow-on jobs for rapid follow-on execution
 - Stores job execution state long-term in a job logging database



CATALYST Implementation

- Primary languages Perl & Java
 - Perl ~12.2K LOC
 - Java ~6.6K LOC
- Defines an Application Programming Interface (API) using externally accessible XML-RPC
 - Allows users to inspect and modify job execution status programmatically in any language (with XML-RPC libraries)
 - Enables development of modular tools to function in concert with CATALYST – backlogging component
- Uses existing AMI LDAP service for user authentication
 - Permissions defined in access control list



Timeline of Events

- **January 2012 – ASDC & DMT workshop**
- **3/1/2013 – PR Tool went “Live”**
- **3/22/2013 – CATALYST Test Readiness Review**
- **5/1/2013 – Implemented automation software Change Review Board**
- **11/19/2013 – Release current CATALYST Server v1.0.4**
- **12/9/2013 – Release current Operator Console v1.0.4**
- **12/19/2014 – CATALYST Operational Readiness Review**
- **3/18/2014 – Rebuild and update Logging DB (reflect DPO post recovery)**
- **3/18/2014 – Release standalone status monitoring & recovery utilities**
- **4/16/2014 – CATALYST Go-Live**



Where Are We Now?



**YOU
ARE
HERE**

- **CATALYST Release 1 (Build 1.0.4)**
 - **Server core functionality implemented**
 - **XML-RPC based API implemented**
 - **Modular framework implemented**
 - **PGE logic encapsulated in handlers**
 - **Beta2_Ed4 SSF Stream (8 PGE Handlers)**
 - **About 85% of Ops Console functionality implemented**



Lessons Learned

- **Industry standard ticketing system – JIRA**
- **Human In The Loop**
 - Can pre-identify most “won’t run” scenarios when reprocessing
 - Forward processing often requires judgment
 - Operator manages Exceptions for off nominal cases
 - Operator can fine tune system load until we observe enough use cases to begin intelligently weighting jobs
- **Change Review Board – Anyone can “request a change”**

CATALYST Operator's Console

The screenshot displays the CATALYST Operator's Console interface. The top menu bar includes 'File', 'CATALYST Server', 'PR Management', 'Tools', and 'Help'. Below the menu, a status bar shows 'Pending PR Ingests: 0', 'CERES Production CATALYST Server', 'Server Version: 1.0.4', 'Cluster Status: [Green]', 'CATALYST Mem Usage: 869M', 'Pending PR Approvals: 0', and 'Pending PR Deletions: 0'.

The main content area is divided into two panes. The left pane, titled 'Production Requests', shows a tree view of requests. The right pane, titled 'Jobs', displays a table of job details.

Production Requests Tree View:

- Production Requests
 - Clouds
 - CER4.1-4.1P6
 - CER4.1-4.1P6 Aqua-MODIS 586-14
 - CER4.1-4.1P6 Terra-MODIS 585-14
 - CER4.1-4.2P4
 - CER4.1-4.2P4 Aqua-MODIS 526-14
 - CER4.1-4.2P4 Terra-MODIS 529-14
 - CER4.1-4.2P5
 - CER4.1-4.2P5 Aqua-MODIS 584-14
 - CER4.1-4.2P5 Terra-MODIS 583-14
 - CER4.1-4.3P3
 - CER4.1-4.3P3 Aqua-MODIS 527-14
 - CER4.1-4.3P3 Terra-MODIS 528-14
 - Inversion
 - CER4.5-6.1P4
 - CER4.5-6.1P4 Terra-FM1-MODIS 530-14
 - CER4.5-6.1P4 Terra-FM2-MODIS 530-14
 - CER4.5-6.1P5
 - CER4.5-6.1P5 Aqua-FM3-MODIS 533-14
 - CER4.5-6.1P5 Aqua-FM4-MODIS 533-14
 - CER4.5-6.2P3
 - CER4.5-6.2P3 Aqua-FM3-MODIS 534-14

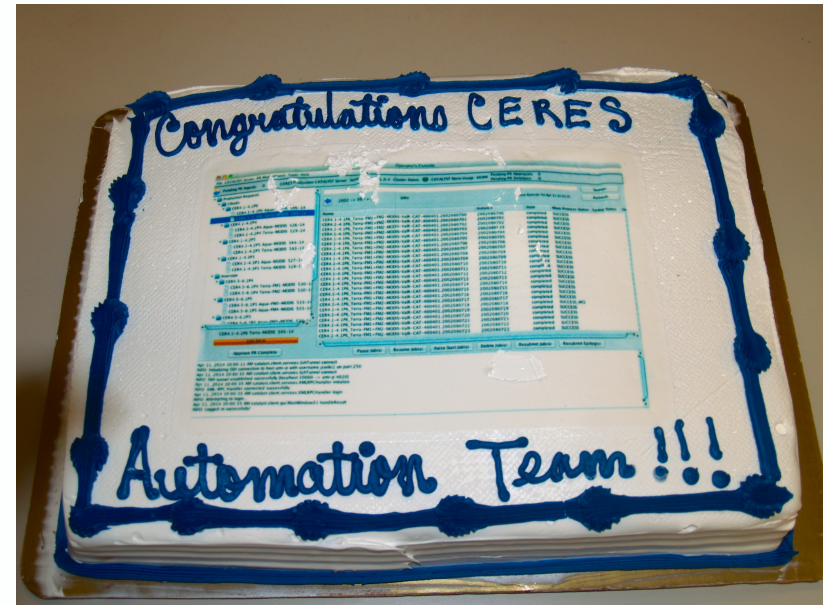
Jobs Table:

Name	Instance	State	Main Process Status	Epilog Status
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080700	2002080700	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080701	2002080701	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080702	2002080702	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080703	2002080703	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080704	2002080704	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080705	2002080705	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080706	2002080706	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080707	2002080707	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080708	2002080708	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080709	2002080709	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080710	2002080710	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080711	2002080711	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080712	2002080712	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080713	2002080713	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080714	2002080714	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080715	2002080715	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080716	2002080716	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080717	2002080717	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080718	2002080718	completed	SUCCESS_WQ	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080719	2002080719	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080720	2002080720	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080721	2002080721	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080722	2002080722	completed	SUCCESS	
CER4.1-4.1P6_Terra-FM1+FM2-MODIS-ValR-CAT-400401.2002080723	2002080723	completed	SUCCESS	

The bottom section of the console displays a log of system events:

```
Apr 11, 2014 10:00:11 AM catalyst.client.services.SshTunnel connect
INFO: Initializing SSH connection to host:ami-p with username:jcwilk1 on port:250
Apr 11, 2014 10:00:35 AM catalyst.client.services.SshTunnel connect
INFO: SSH tunnel established successfully [localhost:10000 -> ami-p:4020]
Apr 11, 2014 10:00:35 AM catalyst.client.services.XMLRPCHandler initialize
INFO: XML-RPC Handler connected successfully.
Apr 11, 2014 10:00:35 AM catalyst.client.services.XMLRPCHandler login
INFO: Attempting to login...
Apr 11, 2014 10:00:35 AM catalyst.client.gui.MainWindow$1 handleResult
INFO: Logged in successfully!
```


CATALYST Go Live!



The Science Directorate at NASA's Langley Research Center



CATALYST Release Schedule

- **Release 1:**
 - Build 1.0.4 – Currently in production
 - Build 1.0.5 - Patch delivery
 - Deliver for testing on 4/24/2014
 - Promote to production on 5/27/2014
 - Build 1.0.6 – Patch delivery
 - Deliver for testing on 7/11/2014
 - Promote to production on 8/24/2014



CATALYST Release Schedule (Cont'd)

- **Release 2: Edition 4 SSF1deg-Stream**
 - Implement Improved log database API
 - Implement improved job inspection API
 - Add Edition 4 Inversion PGE Handlers
 - Add Edition 4 SSF1deg-Hr PGE Handlers
 - Add Edition 4 SSF1deg-Day/Month Handlers
- **Release 3: Edition 4 SYN1deg Stream**
 - Add Edition 4 TSI-> SYNI-> SYN1deg PGE Handlers



CATALYST Release Schedule (Cont'd)

- **Release 4: Edition 3/4 Instrument & ERBElike Stream**
 - Add Edition 3/4 Instrument PGE Handlers
 - Add Edition 3/4 ERBElike PGE Handlers
- **Release 5: Baseline1-QC / Edition1-CV Instrument Stream**
 - Add Edition 1-CV Instrument PGE Handlers



CATALYST Release Schedule (Cont'd)

- **Release 6: Edition 4 CRS and Edition 1 NPP CRS Stream**
 - Add CRS PGE Handlers
- **Release 7: Edition 4 ISCCP-D2Like & Flux-By-Cloud Streams**
 - Add Edition 4 ISCCP-D2Like PGE Handlers
 - Add Edition 4 Flux-By-Cloud PGE Handlers



CATALYST Development & Delivery Schedule Summary View



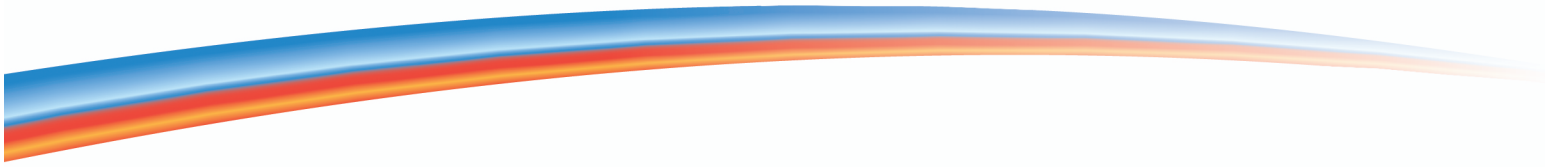
4/22/14

Key Milestones	FY14				FY15			
	1Q'14	2Q'14	3Q'14	4Q'14	1Q'15	2Q'15	3Q'15	4Q'15
Release 1: CATALYST Development & Deployment								
Operational Readiness Review (ORR)		▲ 12/19						
Run & Validate ValRs		4/8	4/10					
Promote to Production			▲ 4/16					
Build 1.0.5								
Development, Internal Testing, and Delivery to CM		3/24	4/24					
Operational Testing		4/25	5/9					
ANGe Outage			5/12	5/16				
Complete Operational Testing/Promote to Production			5/19	5/27				
Build 1.0.6								
Development, Internal Testing, and Delivery to CM			5/28	7/11				
Operational Testing/Promote to Production			7/14	8/4				
Release 2: Edition 4 SSF-1deg Stream								
Development, Internal Testing, and Delivery to CM				8/5	10/15			
Operational Testing/Promote to Production				10/15	11/27			
Release 3: Edition 4 SYN1deg Stream								
Development, Internal Testing, and Delivery to CM					11/27	1/12		
Operational Testing/Promote to Production					1/13	2/3		
Release 4: Edition 3/4 Instrument & ERBElike Stream								
Development, Internal Testing, and Delivery to CM						2/4	3/6	
Operational Testing/Promote to Production						3/6	3/26	
Release 5: Edition 1-CV/Baseline1-QC Instrument Stream								
Development, Internal Testing, and Delivery to CM						3/26	4/27	
Operational Testing/Promote to Production						4/28	5/12	
Release 6: Edition 4/NPP CRS								
Development, Internal Testing, and Delivery to CM							5/13	6/1
Operational Testing/Promote to Production							6/2	6/19
Release 7: Edition 4 ISCCP & Flux-By-Cloud								
Development, Internal Testing, and Delivery to CM							6/22	7/23
Operational Testing/Promote to Production							7/24	8/12

Work Effort

Open Milestone

Completed Milestone



Terra / Aqua Edition 3 & 4 & NPP Processing Approach

Edition 4 Planned Milestones

Product	Science Delivery to DMT	Target Public Release
Ed4 Inversion	ADM November 22, 2013	May 28, 2014
	SOFA January 9, 2014	
Ed4 SSF1deg-Hr	April 21, 2014	June 25, 2014
Ed4 SSF1deg-Day/Month	May 23, 2014	July 15, 2014
Ed4 TSI	June 27, 2014	October 27, 2014
Ed4 SYNI	July 11, 2014	
Ed4 SYN1deg	July 25, 2014	
Ed4 ISCCP-D2like Day/Nit + GEO + MRG	August 29, 2014	March 5, 2015
Ed4 Flux-By-Cloud	October, 31, 2014	December 31, 2014

NPP & Edition 3 Planned Milestones

Product	Science Delivery to DMT	Target Public Release
Ed3 SSF1deg-Day/Month	May 23, 2014	August 21, 2014
NPP Ed1 Clouds	February 14, 2014	September 26, 2014
NPP Ed1 Inversion	May 9, 2014	
NPP Time Varying Gains (Jan 2012 – Feb 2014)	May 23, 2014	June 13, 2013 (BDS & ERBElke)
NPP Ed1 CRS	November 14, 2014	January 14, 2015

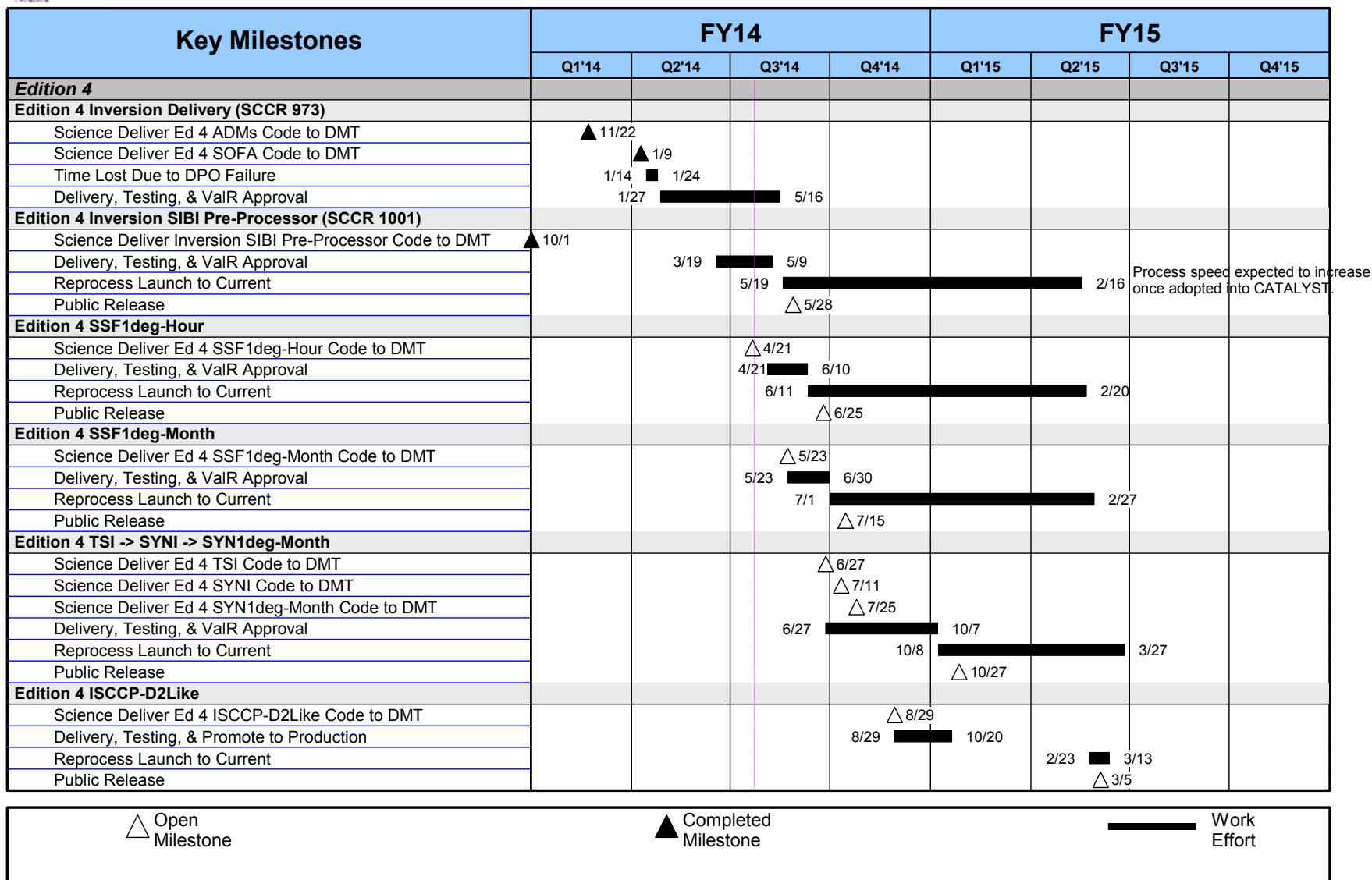


CERES Software Delivery and Data Processing Schedule

Edition 4, NPP, & Edition 3



4/22/14





CERES Software Delivery and Data Processing Schedule Edition 4, NPP, & Edition 3



4/22/14

Key Milestones	FY14				FY15			
	Q1'14	Q2'14	Q3'14	Q4'14	Q1'15	Q2'15	Q3'15	Q4'15
Edition 4 Flux-by-Cloud Type								
Science Deliver Ed 4 Flux-by-Cloud Type Code to DMT					△ 10/31			
Delivery, Testing, & ValR Approval					10/31 ■ 12/18			
Reprocess Launch to Current					12/19 ■ 1/16			
Public Release					△ 12/31			
NPP								
NPP Edition 1 SSF								
Science Deliver Ed 1 Clouds Code to DMT		▲ 2/14						
DMT Prepare Delivery Package Using Ed4 ADM Code		4/7 ■ 5/9						
Delivery, Testing, & ValR Approval		2/18 ■ 7/1						
Reprocess Launch to Current			7/2 ■ 9/18					
Public Release					△ 9/26			
Produce Edition 1 IES with Time Varying Gains								
Instrument Group Deliver NPP Time Varying Gains			△ 5/23					
Reprocess Instrument (via 1.4P3)			5/23 ■ 5/30					
Publicly Release NPP BDS with Time Varying Gains			△ 6/13					
Optional Redelivery of NPP Inversion-Only Code								
Science Deliver Edition 1 SOFA Code to DMT					△ 10/10			
NPP Edition 1 CRS (SCCR 975)								
Science Deliver Ed 1 CRS Code to DMT					△ 11/14			
Delivery, Testing, & ValR Approval					11/14 ■ 1/5			
Reprocess Launch to Current					1/6 ■ 1/12			
Public Release					△ 1/14			
Edition 3								
Edition 3 SSF1-deg-Day/Month								
Science Deliver Ed 3 SSF1deg-Month Code to DMT			△ 5/23					
Delivery, Testing, & ValR Approval			5/23 ■ 8/12					
Reprocess Launch to Current				8/13 ■ 8/26				
Public Release				△ 8/21				

△ Open
Milestone

▲ Completed
Milestone

■ Work
Effort

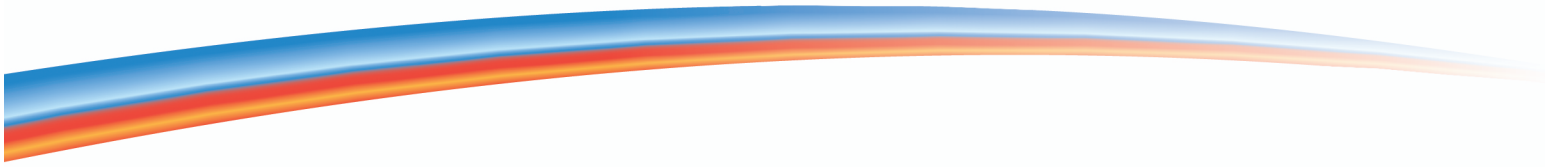
CERES Data Users Statistics

Number of Unique Users By Product

Products	2010	2011	2012	2013	2014
EBAF-TOA	72	146	234	379	207
EBAF-Surface			118	217	112
ESG		14	130	151	36
SYN1deg			61	315	115
SYN1deg-lite	41	126	133		
SSF1deg-lite	46	106	93	138	55
ISCCP-D2like	17	12	45	62	18
SSFlevel2	84	77	138	174	63
BDS	11	9	14	8	1
ES4	59	36	11	17	3
ES8	22	20	18	21	4
ES9	21	12	5	9	1
SFC	31	20	14	6	0
NEWS	31	32	33	19	2
MISR-MODIS	9	4	2	5	3
FLASH_SSF	25	8	15	12	4
FLASH_TISA	17	18	20	11	1

Total Number of Unique Users

2010	2011	2012	2013	2014
486	640	1084	1544	625



Backup Charts